

3 determining whether a sum of the short-term averaged energy and a factor is greater
4 than the long-term averaged energy; and
5 determining that the current audio frame represents silence if the sum is less than the
6 long-term averaged energy, without necessitating a determination of the peak-to-mean
7 likelihood ratio.

1 4. The method of claim 3, upon determining that the sum is greater than the
2 long-term averaged energy and before determining the peak-to-mean likelihood ratio, the
3 method further comprises:
4 determining whether a difference between the long-term averaged energy and the
5 short-term averaged energy is less than a predetermined threshold;
6 determining that the current audio frame represents voice if the difference is greater
7 than the predetermined threshold; and
8 continuing by determining the peak-to-mean likelihood ratio if the difference is less
9 than the predetermined threshold.

1 5. The method of claim 2, wherein the determining of the short-term averaged
2 energy comprises:
3 determining an energy, in decibels, of the current audio frame;
4 determining a short-term averaged energy for a prior audio frame; and
5 conducting a weighted average of the energy of the current audio frame and the short-
6 term averaged energy for the prior audio frame.

1 6. (Twice Amended) A method for enhancing voice activity detection
2 comprising:
3 determining a peak-to-mean likelihood ratio, the determining a peak-to-mean
4 likelihood ratio comprises

5 calculating an averaged peak-to-mean ratio for the current audio frame,
6 determining a maximum averaged peak-to-mean ratio,
7 determining a minimum averaged peak-to-mean ratio,
8 determining a difference between the maximum averaged peak-to-mean ratio
9 and the averaged peak-to-mean ratio for the current audio frame,
10 determining a difference between the maximum averaged peak-to-mean ratio
11 and the minimum averaged peak-to-mean ratio, and
12 conducting a ratio, a denominator of the ratio being the difference between the
13 maximum averaged peak-to-mean ratio and the minimum averaged peak-to-mean
14 ratio, the numerator being the difference between the maximum averaged peak-to-
15 mean ratio and the averaged peak-to-mean ratio; and
16 comparing the peak-to-mean likelihood ratio to a selected threshold to determine
17 whether a current audio frame represents a voice signal.

1 7. (Cancelled)

1 8. (Amended) The communication module of claim 12, wherein the voice
2 activity detector, when executed, controls the processing unit to determine whether a sum of
3 the short-term averaged energy and a predetermined factor is greater than the long-term
4 averaged energy, and to signal that the current audio frame represents silence if the sum is
5 less than the long-term averaged energy.

1 9. The communication module of claim 8, wherein the voice activity detector,
2 when executed, controls the processing unit to determine whether a difference between the
3 long-term averaged energy and the short-term averaged energy is less than a predetermined
4 threshold, and to signal that the current audio frame represents voice if the difference is
5 greater than the predetermined threshold.

1 10. (Cancelled)

1 11. (Amended) The communication module of claim 9, wherein the voice activity
2 detector, when executed, controls the processing unit to determine a peak-to-mean ratio by (i)
3 sampling an analog signal a predetermined number of times to produce a plurality of sampled
4 signals each having a sampled value, (ii) determining a maximum value of the plurality of
5 sampled signals, and (iii) conducting a ratio between an absolute value of the maximum
6 value and a summation of the sampled values for the plurality of sampled signals.

1 12. (Twice Amended) A communication module comprising:
2 a substrate;
3 a processing unit placed on the substrate; and
4 a memory coupled to the processing unit, the memory to contain a voice activity
5 detector which, when executed, controls the processing unit to
6 determine a peak-to-mean likelihood ratio for the current audio frame by (i)
7 monitoring a maximum averaged peak-to-mean ratio and a minimum averaged peak-
8 to-mean ratio, (ii) determining a first result being a difference between the maximum
9 averaged peak-to-mean ratio and the averaged peak-to-mean ratio for the current
10 audio frame, (iii) determining a second result being a difference between the
11 maximum averaged peak-to-mean ratio and the minimum averaged peak-to-mean
12 ratio, and (iv) conducting a ratio between the first result as a numerator and the
13 second result as a denominator; and
14 compare the peak-to-mean likelihood ration to a selected threshold to
15 determine whether the current audio frame represents a voice signal.

1 13. (Twice Amended) A machine readable medium having embodied thereon
2 a computer program for processing by a machine, the computer program comprising:

3 a first routine for determining a normalized peak-to-mean likelihood ratio including
4 (i) a denominator having a value substantially equal to a difference between a maximum
5 averaged peak-to-mean ratio and a minimum averaged peak-to-mean ratio and (ii) a
6 numerator having a value substantially equal to a difference between the maximum averaged
7 peak-to-mean ratio and the averaged peak-to-mean ratio; and
8 a second routine for comparing the peak-to-mean likelihood ratio to a selected
9 threshold to determine whether an audio frame being transmitted represents a voice signal.

1 14. The machine readable medium of claim 13, wherein the computer program
2 further comprising:

3 a third routine for determining a short-term averaged energy for the audio frame, the
4 third routine being executed before the first and second routines; and

5 a fourth routine for determining a long-term averaged energy for the audio frame, the
6 fourth routine being executed before the first and second routines.

1 15. The machine readable medium of claim 14, wherein the computer program
2 further comprising:

3 a fifth routine for determining whether a sum of the short-term averaged energy and a
4 predetermined factor is greater than the long-term averaged energy, the fifth routine being
5 executed before the first and second routines; and

6 a sixth routine for determining whether a difference between the long-term averaged
7 energy and the short-term averaged energy is less than a predetermined threshold, the sixth
8 routine being executed after determining that the sum is greater than the long-term averaged
9 energy and before execution of the first and second routines.

1 16. The machine readable medium of claim 15, wherein the fifth routine
2 determining that the current audio frame represents silence if the sum is less than the long-
3 term averaged energy.

1 17. The machine readable medium of claim 15, wherein the sixth routine
2 determining that the current audio frame represents voice if the difference is greater than the
3 predetermined threshold.

1 18. (Cancelled)

1 20. A method for enhancing voice activity detection comprising:
2 determining a peak-to-mean likelihood ratio including (i) a denominator having a
3 value substantially equal to a difference between a maximum averaged peak-to-mean ratio
4 and a minimum averaged peak-to-mean ratio and (ii) a numerator having a value
5 substantially equal to a difference between the maximum averaged peak-to-mean ratio and
6 the averaged peak-to-mean ratio; and
7 comparing the peak-to-mean likelihood ratio to a selected threshold to determine
8 whether a current audio frame represents a voice signal.

1 21. The method of claim 20, wherein prior to determining the peak-to-mean
2 likelihood ratio, the method further comprises:
3 determining a short-term averaged energy for the current audio frame; and
4 determining a long-term averaged energy for the current audio frame.

1 22. The method of claim 21, wherein after determining the short-term averaged
2 energy and the long-term averaged energy, the method further comprises:
3 determining whether a sum of the short-term averaged energy and a factor is greater
4 than the long-term averaged energy; and

5 determining that the current audio frame represents silence if the sum is less than the
6 long-term averaged energy, without necessitating a determination of the peak-to-mean
7 likelihood ratio.

1 23. The method of claim 22, upon determining that the sum is greater than the
2 long-term averaged energy and before determining the peak-to-mean likelihood ratio, the
3 method further comprises:
4 determining whether a difference between the long-term averaged energy and the
5 short-term averaged energy is less than a predetermined threshold;
6 determining that the current audio frame represents voice if the difference is greater
7 than the predetermined threshold; and
8 continuing by determining the peak-to-mean likelihood ratio if the difference is less
9 than the predetermined threshold.

1 24. The method of claim 21, wherein the determining of the short-term averaged
2 energy comprises:
3 determining an energy, in decibels, of the current audio frame;
4 determining a short-term averaged energy for a prior audio frame; and
5 conducting a weighted average of the energy of the current audio frame and the short-
6 term averaged energy for the prior audio frame.